

# **Deteksi Anomali pada Proses Bisnis dengan Fuzzy Multi Criteria Decision Making dan Decision Tree**

5111100051 Uswatun Hasana Kunio

Dosen Pembimbing I : Prof. Drs. Ec. Ir. Riyanarto Sarno, M.Sc., Ph.D.

Dosen Pembimbing II: Abdul Munif, S.Kom., M.Sc

# Latar Belakang

Banyak terdapat anomali pada proses bisnis suatu perusahaan yang bisa merujuk kepada *fraud*

*Process mining* belum dapat memenuhi kebutuhan pengecekan data di dalam *event logs*.

# Tujuan

Menggunakan model ontologi untuk mendapatkan anomali dengan menarik kesimpulan antara SOP dengan event logs.

Memanfaatkan Fuzzy MCDM TOPSIS dan *Decision Tree* untuk mengetahui kecenderungan aktivitas yang berpotensi terjadi anomali.

# Rumusan Masalah

1. Bagaimana mendapatkan data anomali pada event logs dari proses bisnis yang dijalankan?
2. Bagaimana cara menghitung rating anomali dengan menggunakan Fuzzy Multi Criteria Decision Making for the Technique for Order Preference by Similarity to Ideal Solution (FMCDM TOPSIS)?

# Rumusan Masalah

3. Bagaimana menyimpulkan pola aktivitas dalam event logs yang merupakan pola terjadinya anomali dengan menggunakan Decision Tree?
4. Bagaimana melakukan kategorisasi hasil pendeteksian anomali ke dalam kelas anomali yang sesuai?

# Batasan Masalah

Petri Net

*Event logs*

Simulasi proses bisnis  
aplikasi kredit pada  
bank

Model Ontologi

1200 data

Subjektif dan  
Matematis

Pendeteksian  
anomali

Reasoner Pellet

# Gambaran Umum Sistem

**1**

Mengubah event logs dan SOP ke dalam model ontologi

**2**

Melakukan *conformance checking* terhadap event logs dan SOP

**3**

Hitung nilai rating menggunakan FMCDM TOPSIS

**4**

Mencari label kelas dengan menggunakan rating dan SOP

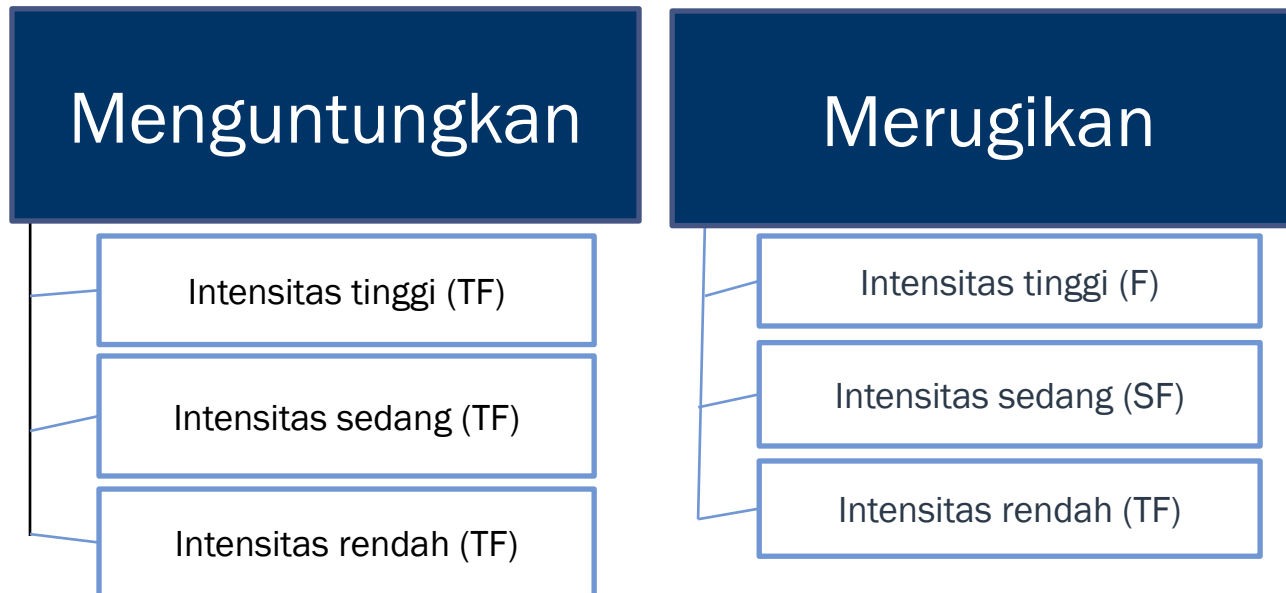
**5**

Memprediksi satu anomali setiap kasus menggunakan Decision Tree

**6**

Hitung akurasi menggunakan Confusion Matrix.

# Anomali



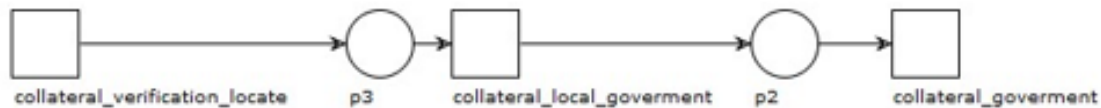
## Jenis Anomali :

*Skip Sequence, Skip Decision, Throughput Time Minimum, Throughput Time Maximum, Wrong Resource, Wrong Duty Sequence, Wrong Duty Decision, Wrong Duty Combine, Wrong Pattern, Wrong Decision.*

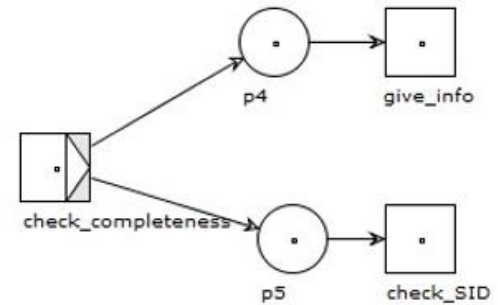


# Anomali

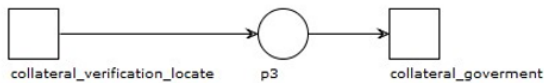
Anomali yang terjadi dilihat dari alur kejadian.



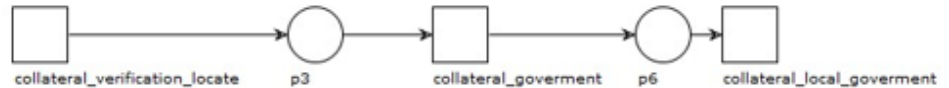
*Aktivitas Sequence*



*Aktivitas Decision*



*Skip Sequence*

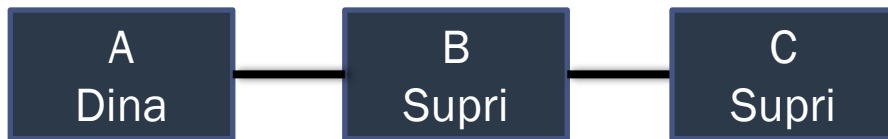


*Wrong Pattern*

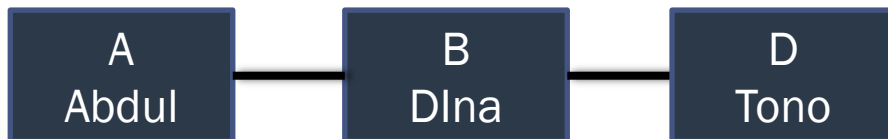
# Anomali

Anomali yang terjadi dilihat dari *role* dan *resource*.

Aktivitas	Originator	Posisi
A	Dina, Dian	Operator
B	Abdul, Dina	Operator
C	Yoga, Supri	Supervisor
D	Tono, Budi	Supervisor
E	Farid	Authorizer
F	Farid, Kasim	Authorizer



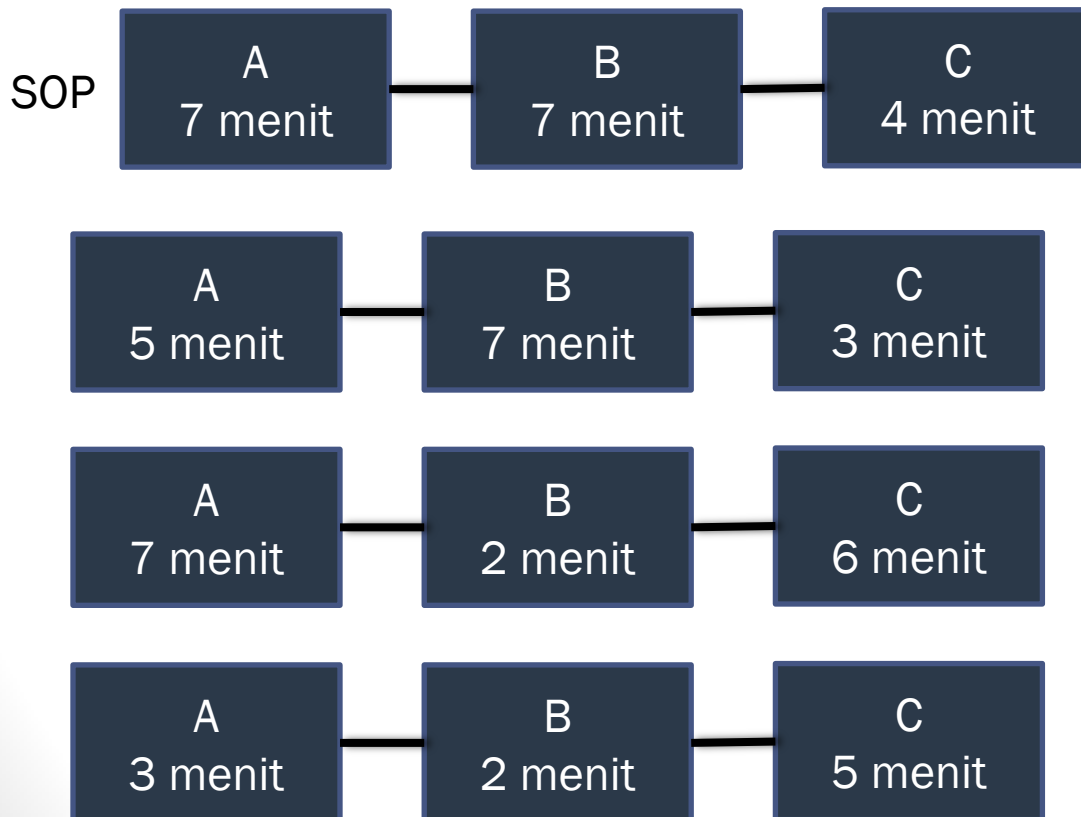
*Wrong Duty*



*Wrong Resource*

# Anomali

Anomali yang terjadi dilihat dari *waktu*.

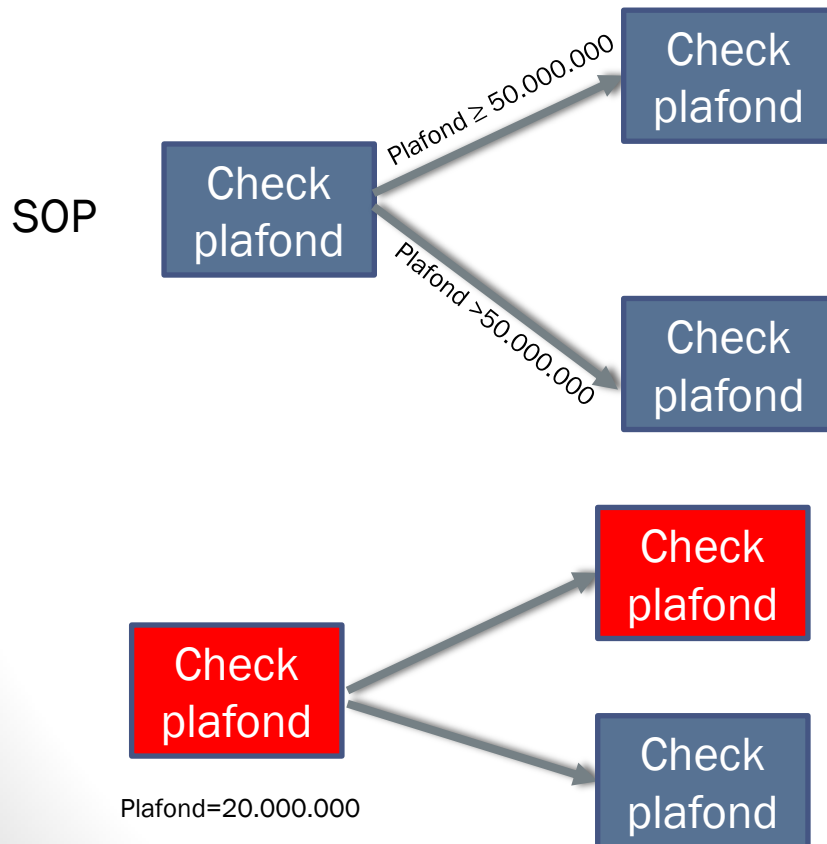


- Cari rata-rata (mean) setiap activity
- Cari selisih antara SOP dengan mean untuk mendapatkan nilai toleransi
- Hitung Confidence Interval untuk mencari titik maksimum dan minimum pelanggaran
- Hitung nilai pelanggaran.



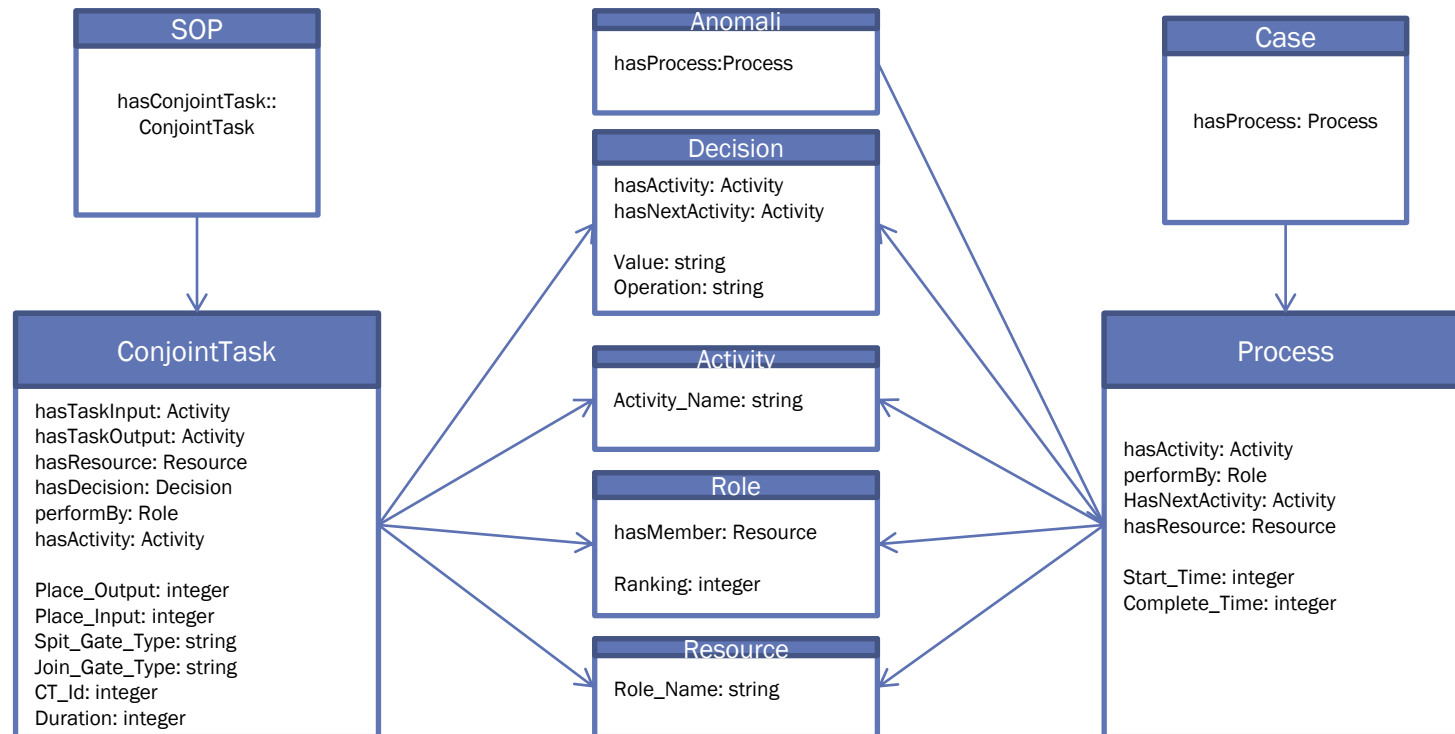
# Anomali

Anomali yang terjadi dilihat dari kondisi kesalahan pengambilan keputusan.



# Model Ontologi

Menurut definisi dari W3C consortium[7] *ontology* adalah suatu istilah yang digunakan untuk menjelaskan dan merepresentasikan sebuah pengetahuan atau *knowledge*



# Rules

- Rules merupakan penarikan kesimpulan dari fakta-fakta yang ada.
  - Menggunakan bahasa SWRL
  - Dijalankan dengan Pellet Reasoner

```
Activity(?a), Case(?c), ConjointTask(?ct), Process(?e),  
SOP(SOP), hasConjointTask(SOP, ?ct), hasProcess(?c,  
?e), hasActivity(?ct, ?a1), hasActivity(?e, ?a2),  
Activity_Name(?a1, ?n1), Activity_Name(?a2,  
?n2), equal(?n1, ?n2),  
hasNextActivity(?e, ?na1), hasTaskOutput(?ct, ?na2),  
Activity_Name(?na1, ?nm1), Activity_Name(?na2,  
?nm2), notEqual(?nm1, ?nm2) ->  
hasCase(WrongPattern, ?e)
```

Contoh rules untuk *wrong pattern*

# Rules

The screenshot displays a Semantic Web editor interface with several panels:

- Individuals: WrongPattern**: A list of individuals including heru, leader\_otorization, loan\_decision, loan\_drowdone, loan\_reject, make\_validation, Nandes, nandes, Operator, plafond\_estimation, plafond\_estimation\_adm, receive\_application, SkipDecision, SkipSequence, SOP, Supervisor, Susi, Tegar, tegar, ThroughputTimeMax, ThroughputTimeMin, Tiara, WrongDecision, WrongDutyCombine, WrongDutyDecision, WrongDutySequence, WrongPattern (selected), and WrongResource.
- Rules**: Contains a single rule: `Case(?c), SOP(SOP), hasActivity(?ct, ?a1), hasActivity(?e, ?a1), hasConjointTask(SOP, ?ct), hasNextActivity(?e, ?na1), hasProcess(?c, ?e), hasTaskOutput(?ct, ?na2), Activity_Name(?na1, ?nm1), Activity_Name(?na2, ?nm2), notEqual(?nm1, ?nm2) -> hasCase(WrongPattern, ?e)`
- Annotations: WrongPattern**: Shows an annotation `PN_sg` for the `WrongPattern` class.
- Description: WrongPattern**: Shows the class `Anomali` as a type.
- Property assertions: WrongPattern**: Lists object property assertions for `hasCase` with values: Event\_81\_9, Event\_81\_12, Event\_81\_5, Event\_81\_2, Event\_81\_17, Event\_81\_14, Event\_81\_3, and Event\_81\_18.

# METODOLOGI



# Metodologi

- Conformance Checking
- Perhitungan *Rating* menggunakan metode FMCDM TOPSIS
- *Data Training* dan *Data Testing* menggunakan Decision Tree

# Pairwise matrix



1

Bobot kepentingan antar kriteria

	Skip	TT	WDec
Skip	1	1	2
TT	1	1	1/2
Wdec	1/2	2	1

[Lihat dokumen lebih rinci](#)

# Fuzzy Pairwise matrix

2

Intensity of Importance	Definition	Fuzzy Triangular Scale					
		Scale			Reciprocal Scale		
		low	mid	up	low	mid	up
1	Equally important	1	1	1	1	1	1
3	Weakly important	2	3	4	0,25	0,33	0,5
5	Fairly important	4	5	6	0,17	0,2	0,25
7	Strongly important	6	7	8	0,13	0,143	0,167
9	Absolutely important	9	9	9	0,11	0,11	0,11
2	Intermediate values	1	2	3	0,33	0,5	1
4	Intermediate values	3	4	5	0,2	0,25	0,33
6	Intermediate values	5	6	7	0,143	0,167	0,2
8	Intermediate values	7	8	9	0,11	0,125	0,143

	Skip	TT	WDec
Skip	1,1,1	1,1,1	1,2,3
TT	1,1,1	1,1,1	0.33,0.5,1
Wdec	0.33,0.5,1	1,2,3	1,1,1

[Lihat dokumen lebih rinci](#)

$$r_i = \left( \prod_{j=1}^n d_{ij} \right)^{1/n}$$

# Geometric mean



2

Kriteria	Geometric mean		
	low	Mid	high
Skip	$(1*1*1)^{1/3} = 1$	$(1*1*2)^{1/3} = 1.23$	$(1*1*3)^{1/3} = 1.44$
TT	$(1*1*0.33)^{1/3} = 0.11$	$(1*1*0.5)^{1/3} = 0,8$	$(1*1*1)^{1/3} = 1$
WDec	$(0.33*1*1)^{1/3} = 0.11$	$(0.5*2*1)^{1/3} = 1$	$(1*3*1)^{1/3} = 1.44$
Total	1.22	3,03	3,88
reverse	$(1,22)^{-1}$	$(3,03)^{-1}$	$3,88^{-1}$
Increasing order	0,26	0,33	0,82

$$w_i = r_i \otimes (r_1 \oplus r_2 \oplus \dots \oplus r_n)^{-1}$$

# Bobot Kriteria Fuzzy

2

Kriteria	Geometric mean		
	low	Mid	high
Skip	$1 * 0.26 = 0,26$	$1.23 * 0.33 = 0,41$	$1.44 * 0.82 = 1,18$
TT	$0.11 * 0,26 = 0,0286$	$0,8 * 0,33 = 0,264$	$1 * 0,82 = 0,82$
WDec	$0.11 * 0,26 = 0,0286$	$1 * 0,33 = 0,33$	$1.44 * 0,82 = 1,18$

# FMCDM TOPSIS

*FMCDM Technique for Order Preference by Similarity to Ideal Solution*

Konsep:

tiap case harus memiliki **jarak terdekat dengan solusi ideal-positif** dan **jarak terjauh dari solusi ideal-negatif**

# FMCDM TOPSIS

## Fuzzyfikasi Matriks Data Pelanggaran

### Matriks Data Pelanggaran

SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec
0	0	0	0	0	0	0	0,11	0	0
0	0	0	0	0	0,3	0	0	0	0
0	0	0	0	0	0,7	0	0	0	0

### Fungsi Membership Derajat Pelanggaran

range value	Keterangan	Atribut
$0 < x < 0,1$	VERY WEAK	VW
$0,1 < x < 0,2$	BETWEEN VERY WEAK AND WEAK	BVW&W
$0,2 < x < 0,3$	WEAK	W
$0,3 < x < 0,4$	BETWEEN WEAK AND FAIR	BW&F
$0,4 < x < 0,5$	FAIR	F
		BF&S
		S
		RONG
		BS&VS
		VS

### Matriks Data Pelanggaran

SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec
0	0	0	0	0	0	0	BVW&W	0	0
0	0	0	0	0	BW&F	BS&VS	0	0	0
0	0	0	0	0	BS&VS	0	S	0	0

# FMCDM TOPSIS

## Fuzzyfikasi Matriks Data Pelanggaran

### Matriks Data Pelanggaran

SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec
0	0	0	0	0	0	0	BVW&W	0	0
0	0	0	0	0	BW&F	BS&VS	0	0	0
0	0	0	0	0	BS&VS	0	S	0	0

: Pelanggaran

Keterangan	Atribut	lower	middle1	middle2	upper
VERY WEAK	VW	0,0	0,0	0,1	0,2
BETWEEN VERY WEAK AND WEAK	BVW&W	0,0	0,1	0,2	0,3
WEAK	W	0,1	0,2	0,3	0,4
BETWEEN WEAK AND FAIR	BW&F	0,2	0,3	0,4	0,5
FAIR	F	0,3	0,4	0,5	0,6
BETWEEN FAIR AND STRONG	BF&S	0,4	0,5	0,6	0,7
STRONG	S	0,5	0,6	0,7	0,8
BETWEEN STRONG AND VERY STRONG	BS&VS	0,6	0,7	0,8	1,0
VERY STRONG	VS	0,7	0,8	1,0	1,0

### Matriks Data Pelanggaran

SS	SD	Tmin		Tmax		WR	WDs		WDd		WDc		WP	Wdec			
0	0	SS	SD	Tmin	Tmax	WR	WDs		WDd		WDc		WP	Wdec			
0	0	0	0	SS	SD	Tmin	Tmax	WR	WDs		WDd		WDc		WP	Wdec	
0	0	0	0	0	SS	SD	Tmin	Tmax	WR	WDs		WDd		WDc		WP	Wdec
			0	0	0	0	0	0	0	0	0	0	0,3	0	0		
				0	0	0	0	0	0	0,5	1	0	0	0	0		
					0	0	0	0	0	1	0	0,8	0	0	0		



## Weighted Matrix

## Matriks Data Pelanggaran

SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec					
0	0	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec			
0	0	0	0	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec	
0	0	0	0	0	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec
		0	0	0	0	0	0	0	0	0	0	0	0	0
			0	0	0	0	0	0	0	0	0	0	0	0
				0	0	0	0	0	0	0	0	0	0	0
					0	0	0	0	0	0	0	0	0	0

Bobot Kepentingan K

pelanggaran	lower	middle1	middle2
skip sequence	0,0288	0,0904	0,1520
skip decision	0,0501	0,1809	0,3117

## Bobot Kepentingan Kriteria

pelanggaran				lower	middle1	middle2	upper
skip sequence				0,0288	0,0904	0,0904	0,3103
skip decision				0,0501	0,1809	0,1809	0,5386
throughput time min				0,0452	0,0838	0,0838	0,1542
throughput time max				0,0452	0,0838	0,0838	0,1542
wrong resource				0,0500	0,0990	0,0990	0,2290
				69	0,0227	0,0227	0,0914
				00	0,0752	0,0752	0,2636
				09	0,0414	0,0414	0,1661
WP	Wdec	bobot		80	0,0700	0,0700	0,1580
				20	0,2540	0,2540	0,5080

## Matriks Data Pelanggaran

SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec	bobot		80	0,07
0	0											20	0,25
	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec	bobot		
0	0	0	0	0	0	0	0	0,1*0,041	0	0	0,041		
0	0	0	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec	bobot
	0	0	0	0	0	0	0	0	0,2*0,041	0	0	0,0082	
		0	SS	SD	Tmin	Tmax	WR	WDs	WDd	WDc	WP	Wdec	bobot
		0	0	0	0	0	0	0	0,3*0,09	0	0	0,027	
		0	0	0	0	0	0,5*0,05	1*0,14	0	0	0	0,165	
			0	0	0	0	0	1*0,05	0	0,8	0	0	0,85

## Menentukan Solusi Ideal Positif dan Ideal Negatif

# FMCDM TOPSIS

## *Separation Measure dari Solusi Ideal Positif dan Ideal Negatif*

*Difference Measure dari  
Solusi Ideal Positif:*

$$D_{ij}^* = |M(v_{ij}) - M(v_{ij})^*| \quad (2.6)$$

*Separation Measure dari  
Solusi Ideal Positif:*

$$S_i^* = \sum_{j=1}^n D_{ij}^*, \quad i = 1, \dots, m \quad (2.7)$$

*Difference Measure dari  
Solusi Ideal Negatif:*

$$D_{ij}^- = |M(v_{ij}) - M(v_{ij})^-| \quad (2.7)$$

*Separation Measure dari  
Solusi Ideal Negatif:*

$$S_i^- = \sum_{j=1}^n D_{ij}^-, \quad i = 1, \dots, m \quad (2.7)$$

# FMCDM TOPSIS

*Separation Measure* dari Solusi Ideal Positif dan Ideal Negatif

Jarak Terhadap A-												
Case	Skip		Throughput Time		Wrong Resource	Wrong Duty			Wrong Pattern	Wrong Decisio	S-	
	Sequence	Decision	Min	Max		Sequence	Decision	Combine				
1	0	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	0	0	0	0	0	
6	0	0	0	0	0	0	0	0,0182	0	0	0,0182	
7	0	0	0	0	0	0	0	0	0	0	0	

Jarak Terhadap A*											
Case	Skip		Throughput Time		Wrong Resource	Wrong Duty			Wrong Pattern	Wrong Decisio	S*
	Sequence	Decision	Min	Max		Sequence	Decision	Combine			
1	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496
2	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496
3	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496
4	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496
5	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496
6	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,0528	0,08449	0	0,7314
7	0,1389	0	0,0896	0,089645	0,120346	0,039264	0,116335	0,071	0,08449	0	0,7496

# FMCDM TOPSIS

## Kedekatan Relatif dengan Solusi Ideal

$$\dot{c}_i^* = \frac{s_i^-}{(s_i^* + s_i^-)} \quad (2.6)$$

Case	Ci
1	0
2	0
3	0
4	0
5	0
6	0,0243
7	0
8	0
9	0
10	0
11	0
12	0,175
13	0,1239

# Penentuan *Label* Data

Mengikuti aturan verifikasi:

1. Kasus yang terdapat *Skip* dan *Wrong Decision* pasti mengandung anomali.
2. Kasus yang terdapat *Wrong Resource* dan *Wrong Duty* secara bersamaan pasti mengandung anomali.

Mempertimbangkan rating  
Treshold:

Tipe pelanggaran	Range nilai rating	Nilai Pelanggaran
Tidak fraud	$0 \leq x \leq 0,1$	1
	$0,1 < x < 0,2$	$(0,2 - \text{rating})/0,1$
Semi fraud	$0,1 < x < 0,2$	$(\text{rating}-0,1)/0,1$
	$0,2 \leq x \leq 0,4$	1
Fraud	$0,4 < x < 0,5$	$(0,5-\text{rating})/0,1$
	$0,4 < x < 0,5$	$(\text{bobot}-0,4)/0,1$
	$x \geq 0,5$	1

# Penentuan *Label* Data

[illegible]

# Decision Tree

Matriks Data Pelanggaran

Case	Skip		Throughput Time		Wrong Resource	Wrong Duty			Wrong Pattern	Wrong Decision	Rating	Label
	Sequence	Decision	Min	Max		Sequence	Decision	Combine				
1	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
2	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
3	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
4	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
5	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
6	0	0	0	0	0	0	0	0,11	0	0	0,02428	TIDAK FRAUD
7	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
8	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
9	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD
10	0	0	0	0	0	0	0	0	0	0	0	TIDAK FRAUD

Tabel 3.11. Tipe Pelanggaran

Jumlah Pelanggaran / Maksimal Pelanggaran	Tipe Pelanggaran
$x \leq 0.33$	Low
$0.33 < x \leq 0.67$	Medium
$0.67 < x$	High



# Decision Tree

[illegible]

# Decision Tree

## Memilih atribut *root*

Hitung Entropi Tiap Atribut:

$$Entropy(S) = - \sum_{i=1}^n p_i(c) * \text{Log}_2 p_i(c)$$

Hitung Information Gain Tiap Atribut:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Information Gain terbesar dijadikan sebagai root

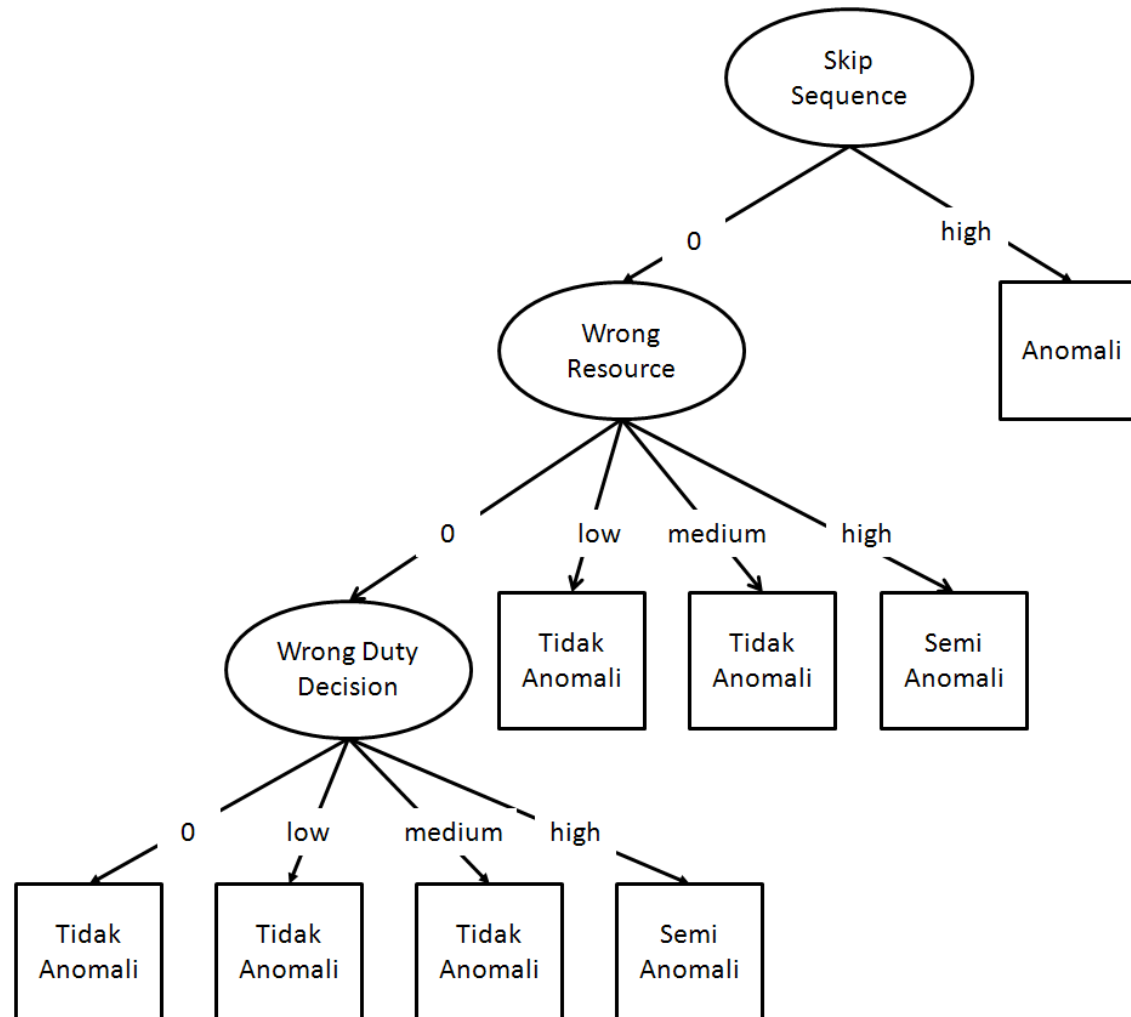
# Decision Tree

Memilih atribut *root*

		Tidak Fraud	Semi Fraud	Fraud	Total	E
W Resource	low	618	20	194	832	0,9377164
	mid	36	4	41	81	1,2314928
	high	42	6	39	87	1,2921643
		Total		1000		
		Entropy(Fraud)		0,99234925		
		Gain		0,035		
		Tidak Fraud	Semi Fraud	Fraud	Total	E
W Duty Sequence	low	613	28	201	842	0,9900162
	mid	50	1	41	92	1,0686441
	high	33	1	32	66	1,0979546
		Total		1000		
		Entropy(Fraud)		1,004373864		
		Gain		0,023		
		Tidak Fraud	Semi Fraud	Fraud	Total	E
W Duty Decision	low	679	29	264	972	1,0234377
	mid	8	0	4	12	0,9182958
	high	9	1	6	16	1,2475562
		Total		1000		
		Entropy(Fraud)		1,025761895		
		Gain		0,002		

# Decision Tree

## *Tree Hasil Data Training*



# Decision Tree

## *Data Testing*

### Confusion Matrix

	A	SA	TA
A	5	0	0
SA	0	3	2
TA	0	2	188

Akurasi : 98%

### Keterangan

A : Anomali

SA : Semi Anomali

TA : Tidak Anomali

# Kesimpulan

- Ada keterbatasan pada SWRL Rules dengan Pellet Reasoner, yaitu tidak dapat menghitung jumlah.
- Deteksi anomali dengan menggunakan FMCDM TOPSIS dan *Decision Tree* mendapat akurasi sebesar 98%

# Saran

- Data yang ada akan lebih baik jika *real* karena persebaran data akan lebih valid. Sehingga berpengaruh saat mengklasifikasikan data.
- Menambah jumlah data latih untuk mengurangi jumlah FP dan FN.
- Klasifikasi data kontinu menjadi diskrit untuk data latih akan lebih baik jika menggunakan metode untuk mengurangi jumlah FP dan FN.

**SEKIAN DAN TERIMA KASIH**